

# ABSTRACT OF REFERENCE 1

(11)Publication number : 11-086448

(43)Date of publication of application : 30.03.1999

---

(51)Int.Cl. G11B 20/10 G10L 3/00 // H03M 1/00

---

(21)Application number : 09-249220 (71)Applicant : NIPPON HOSO KYOKAI

<NHK>

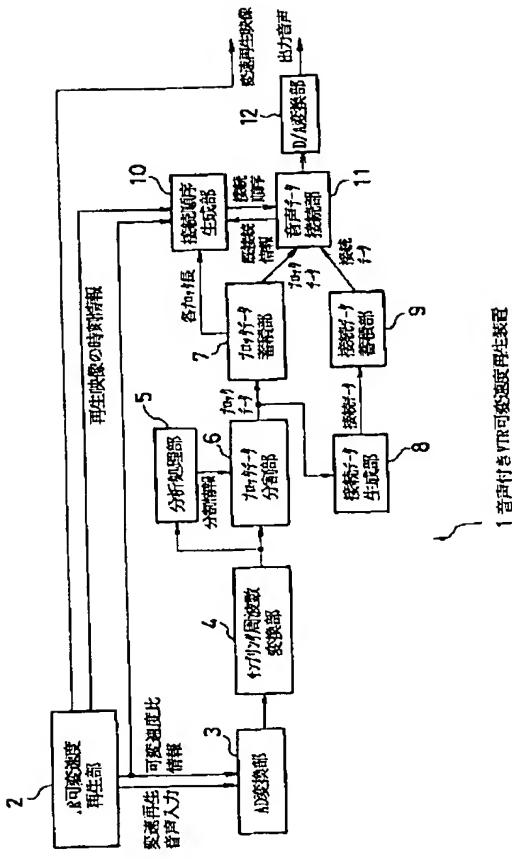
(22)Date of filing : 12.09.1997 (72)Inventor : IMAI ATSUSHI

SEIYAMA NOBUMASA

TSUGI TORU

---

(54) VOICE PROCESSING METHOD, VOICE PROCESSING DEVICE AND  
RECORDING/REPRODUCING DEVICE



(57)Abstract:

PROBLEM TO BE SOLVED: To restore the tone of reproduced voice in the condition that the voice is synchronized with images with a high precision while the clearness of voice is kept by taking in the tone reproduced at a variable speed within the range of variable speed practically usable.

SOLUTION: When voice signals are reproduced at a speed different from that at the time of

recording, the presentation timing of voice is synchronized to the image signals which are reproduced at a variable speed using a VTR variable speed reproducing section 2 and high quality voice whose tone does not vary is outputted by making the sampling frequency  $f_1(\text{Hz})$  at the time of A/D conversion match with the sampling frequency  $f_0(\text{Hz})$  at the time of D/A conversion by a sampling frequency conversion section 4 and at the same time by dividing the voice data to block units having a specified time width, based on the variable speed ratio  $r$  and the attribute of voice, and by successively operating interpolation or thinning.

(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平11-86448

(43)公開日 平成11年(1999)3月30日

(51)Int.Cl.<sup>6</sup>  
G 1 1 B 20/10  
G 1 0 L 3/00  
// H 0 3 M 1/00

識別記号  
3 2 1

F I  
G 1 1 B 20/10  
G 1 0 L 3/00  
H 0 3 M 1/00

3 2 1 Z  
A

審査請求 有 請求項の数8 O.L (全14頁)

(21)出願番号 特願平9-249220

(22)出願日 平成9年(1997)9月12日

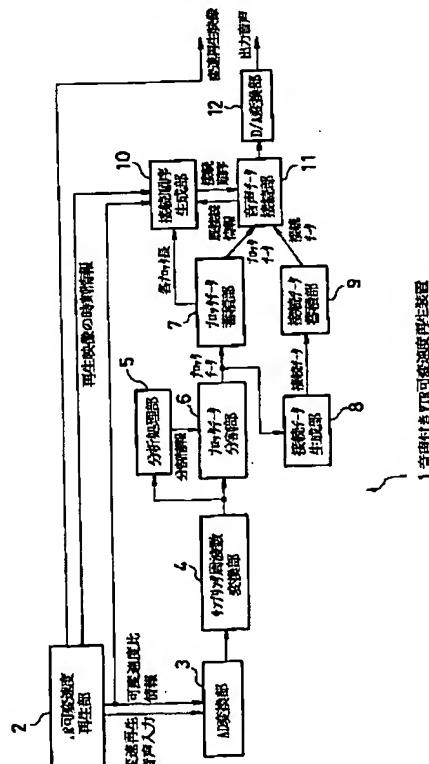
(71)出願人 000004352  
日本放送協会  
東京都渋谷区神南2丁目2番1号  
(72)発明者 今井 篤  
東京都世田谷区砧一丁目10番11号 日本放  
送協会放送技術研究所内  
(72)発明者 清山 信正  
東京都世田谷区砧一丁目10番11号 日本放  
送協会放送技術研究所内  
(72)発明者 都木 徹  
東京都世田谷区砧一丁目10番11号 日本放  
送協会放送技術研究所内  
(74)代理人 弁理士 三好 秀和 (外8名)

(54)【発明の名称】 音声処理方法、音声処理装置および記録再生装置

(57)【要約】

【課題】 可変速度の実用的な使用範囲内で、可変速度再生された音声を取り込んで、音声の明瞭さを保ちながら、映像と高精度に同期させた状態で、再生された音声の音程を元に戻す。

【解決手段】 録音時と異なる速度で音声信号を再生したとき、サンプリング周波数変換部4によってA/D変換時のサンプリング周波数 $f_i$  (Hz)と、D/A変換時のサンプリング周波数 $f_o$  (Hz)とをマッチングさせながら、VTR可変速度再生部2の可変速度比 $r$ と、音声の属性とに基づき、音声データを所定の時間幅を有するブロック単位に分割して逐次補間、または間引き操作することにより、VTR可変速度再生部2で可変速度再生された映像信号に、音声の提示タイミングを同期させながら、音程が変化しない高品質な音声を出力させる。



## 【特許請求の範囲】

【請求項1】 記録媒体に音声信号を記録したときの速度と異なる再生速度で、前記記録媒体上の音声信号を再生する際に、再生された音声信号の音程を調整する音声処理方法において、

正規の再生速度に対する可変速度比  $r$  で、前記記録媒体から再生された音声信号を A/D 変換して音声データにするときのサンプリング周波数  $f_i$  と、前記音声データをデジタル信号処理した後に D/A 変換して音声信号にするときのサンプリング周波数  $f_o$  と、前記速度比  $r$  とが所定の条件を満たすように、各サンプリング周波数  $f_i$ 、 $f_o$  を設定する、

ことを特徴とする音声処理方法。

【請求項2】 請求項1に記載の音声処理方法において、

正規の再生速度に対する可変速度比  $r$  で、前記記録媒体から再生された音声信号をサンプリング周波数  $f_i$  で A/D 変換して得られた音声データをデジタル信号処理する際、

前記音声データの属性を分析処理し、次にこの分析処理で得られた情報に基づき、前記音声データを所定の時間幅を有するブロック単位に分割し、ブロック単位で逐次、補間処理または間引き処理を行なって、データ長を逐次、 $1/r$  倍にする一連の処理を行なって出力音声データを作成した後、

前記サンプリング周波数  $f_o$  で前記出力音声データを D/A 変換して音声信号を生成する、

ことを特徴とする音声処理方法。

【請求項3】 請求項1、または2のいずれかに記載の音声処理方法において、

前記所定の条件として、前記各サンプリング周波数  $f_i$ 、 $f_o$  と、前記速度比  $r$  とが、

$$f_i / f_o = r$$

を満たすときには、

$$f_i = r f_o$$

となるサンプリング周波数  $f_i$  を使用し、

また前記各サンプリング周波数  $f_i$ 、 $f_o$  と、前記可変速度比  $r$  とが、

$$f_i / f_o \neq r$$

であるときには、サンプリング周波数変換係数  $c$  を含む

$$f_i = r f_o / c$$

となるサンプリング周波数  $f_i$  を使用する、ことを特徴とする音声処理方法。

【請求項4】 請求項3に記載の音声処理方法において、

正規の再生速度に対する可変速度比  $r$  で、前記記録媒体から再生された音声信号をサンプリング周波数  $f_i$  で A/D 変換して得られた音声データをデジタル信号処理する際、

前記記録媒体から読み出される音声信号の時刻情報と前

記一連の処理を経て得られた出力音声データの時刻情報を監視し、前記音声信号と前記出力音声データとの間に発生する時間のずれを補正する、

ことを特徴とする音声処理方法。

【請求項5】 記録媒体に音声信号を記録したときの速度と異なる再生速度で、前記記録媒体上の音声信号を再生する際に、再生された音声信号の音程を調整する音声処理装置において、

前記記録媒体から再生された音声信号をサンプリング周波数  $f_i$  で A/D 変換する A/D 変換手段と、

A/D 変換して得られた音声データに対して属性に基づく分析処理を行なう分析処理手段と、

この分析処理で得られた情報に基づき、前記音声データを所定の時間幅を有するブロック単位に分割するプロックデータ分割手段と、

分割されたブロック単位で逐次、補間処理または間引き処理を行なって、データ長を  $1/r$  倍にする一連の処理を行なって出力音声データを作成する出力音声データ作成手段と、

20 作成された出力音声データをサンプリング周波数  $f_o$  で D/A 変換して映像と同期した音声を出力する音声出力手段と、

前記 A/D 変換するときのサンプリング周波数  $f_i$  と、前記音声データを D/A 変換して音声信号にするときのサンプリング周波数  $f_o$  と、前記速度比  $r$  とが所定の条件を満たすように、各サンプリング周波数  $f_i$ 、 $f_o$  を設定するサンプリング周波数設定手段と、

を具備することを特徴とする記録再生装置。

【請求項6】 請求項5に記載の記録再生装置において、

30 前記サンプリング周波数設定手段で設定される所定の条件として、前記各サンプリング周波数  $f_i$ 、 $f_o$  と、前記速度比  $r$  とが、

$$f_i / f_o = r$$

を満たすときには、

$$f_i = r f_o$$

となるサンプリング周波数  $f_i$  を使用し、また前記各サンプリング周波数  $f_i$ 、 $f_o$  と、前記可変速度比  $r$  とが、

$$f_i / f_o \neq r$$

であるときには、サンプリング周波数変換係数  $c$  を含む

$$f_i = r f_o / c$$

となるサンプリング周波数  $f_i$  を使用することを特徴とする記録再生装置。

【請求項7】 請求項5、または6のいずれかに記載の記録再生装置において、

前記出力音声データ作成手段は、前記記録媒体から読み出される音声信号の時刻情報と前記一連の処理を経て得られた出力音声データの時刻情報を監視し、前記音声信号と前記出力音声データとの間に発生する時間のずれ

を補正する手段を含むことを特徴とする記録再生装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、業務用VTR装置、6mmテープレコーダ装置など、再生速度に比例して音程が変化した音声を出力する記録再生装置において、記録された音声を任意の速度で再生する際に、再生速度に対応した正規の音程を持つ音声を出力する音声処理方法および記録再生装置に係わり、特にVTR装置に適用するとき、好みの再生速度に即応し、かつ映像に同期した高品質な音声を出力する音声処理方法および記録再生装置に関する。

【0002】【発明の概要】本発明は、業務用VTR装置などにおいて、スローモーション映像などを収録時と異なる速さで再生する場合、再生速度に比例して音程が変化した音声が出力されるが、この音程が変化した音声を任意のサンプリング周波数  $f_i$  でA/D変換して得られた信号を正規の再生速度に対する可変速度比  $r$  の逆倍数  $1/r$  のサンプリング周波数  $f_i/r$  (Hz) のデジタル信号に変換するとともに、音声のタイミングを映像に同期させるため、サンプリング周波数  $f_i/r$  (Hz) にサンプリング変換した音声データを基本周期またはフレーム単位など、所定の時間幅を有するブロック単位に分割して蓄積し、VTR装置からの入力データ時間長に等しくなるように、逐次、ブロック単位での補間、または間引き操作を行なうことにより、VTR装置を任意の速度で再生する際、その速度の可変タイミングに即応し、かつ映像に同期して音程が変化しない音声を出力する音声処理方法と、この音声処理方法を実現する記録再生装置とを提供するものである。

【0003】

【従来の技術】一般に、磁気テープなどのアナログ記録メディアに記録された音声を任意の速度で再生する場合には、その再生速度に比例して音程も変化する。

【0004】また、業務用VTR装置については、近年、映像・音声共にデジタル方式による記録が主流であるが、音声出力については、編集作業の効率化のため、アナログ方式のVTR装置と同様に、再生速度に比例して音程が変化した音声を出力する。

【0005】

【発明が解決しようとする課題】ところで、VTR装置については、従来から映像の可変速度再生を行なっているものの、可変速度再生の映像に同期しながら、音程を正規の高さに補正するという機能を有するものは、これまでになかった。

【0006】また、音程を補正する手段に限定すれば、カラオケ装置などで使われているピッチシフタ（ピッチチェンジャー）が知られているが、例えばこれを前記VTR装置の音程補正に用いると、対象が人の声であるとき、あるいはその補正範囲が大きいとき、例えば  $1/2$

倍速再生のように、1オクターブ下がった音声を元に戻すときなどには、補正された音声に多くの歪感やエコー感を伴うなど、音質に問題があった。

【0007】これらのことから分かるように、音程の変化を補正する際、放送品質を考慮した場合には、これまで適当な装置が存在せず、放送などで使用するVTR装置をスロー再生させると、意図的に音程の変化した非現実的な音声を用いる場合を除き、映像の再生と同時に、再生された音声を使うことができないという問題があつた。

【0008】本発明は上記の事情に鑑み、 $1/2$ 倍速再生または2倍再生したときのように、可変速度の実用的な使用範囲内で、可変速度再生された、音程が変化した音声を取り込んで、音声の明瞭さを保ちながら、映像と高精度に同期させた状態で、 $\pm 1$ オクターブ程度の範囲で、音程を元に戻すことができる音声処理方法および記録再生装置を提供することを目的としている。

【0009】

【課題を解決するための手段】上記の目的を達成するために本発明は、請求項1では、記録媒体に音声信号を記録したときの速度と異なる再生速度で、前記記録媒体上の音声信号を再生する際に、再生された音声信号の音程を調整する音声処理方法において、正規の再生速度に対する可変速度比  $r$  で、前記記録媒体から再生された音声信号をA/D変換して音声データにするときのサンプリング周波数  $f_i$  と、前記音声データをデジタル信号処理した後にD/A変換して音声信号にするときのサンプリング周波数  $f_o$  と、前記速度比  $r$  とが所定の条件を満たすように、各サンプリング周波数  $f_i$  、  $f_o$  を設定することを特徴としている。

【0010】また、請求項2では、請求項1に記載の音声処理方法において、正規の再生速度に対する可変速度比  $r$  で、前記記録媒体から再生された音声信号をサンプリング周波数  $f_i$  でA/D変換して得られた音声データをデジタル信号処理する際、前記音声データの属性を分析処理し、次にこの分析処理で得られた情報に基づき、前記音声データを所定の時間幅を有するブロック単位に分割し、ブロック単位で逐次、補間処理または間引き処理を行なって、データ長を逐次、 $1/r$  倍にする一連の処理を行なって出力音声データを作成した後、前記サンプリング周波数  $f_o$  で前記出力音声データをD/A変換して音声信号を生成することを特徴としている。

【0011】さらに、請求項3では、請求項1、または2のいずれかに記載の音声処理方法において、前記所定の条件として、前記各サンプリング周波数  $f_i$  、  $f_o$  と、前記速度比  $r$  とが、  

$$f_i/f_o = r$$
を満たすときには、  

$$f_i = r f_o$$
となるサンプリング周波数  $f_i$  を使用し、また前記各サ

ンプリング周波数  $f_i$  、  $f_o$  と、前記可変速度比  $r$  とが、

$$f_i / f_o = r$$

であるときには、サンプリング周波数変換係数  $c$  を含む  $f_i = r f_o / c$

となるサンプリング周波数  $f_i$  を使用することを特徴としている。

【0012】さらに、請求項4では、請求項3に記載の音声処理方法において、正規の再生速度に対する可変速度比  $r$  で、前記記録媒体から再生された音声信号をサンプリング周波数  $f_i$  でA/D変換して得られた音声データをデジタル信号処理する際、前記記録媒体から読み出される音声信号の時刻情報と前記一連の処理を経て得られた出力音声データの時刻情報を監視し、前記音声信号と前記出力音声データとの間に発生する時間のずれを補正することを特徴としている。

【0013】さらに、請求項5の記録再生装置では、記録媒体に音声信号を記録したときの速度と異なる再生速度で、前記記録媒体上の音声信号を再生する際に、再生された音声信号の音程を調整する音声処理装置において、前記記録媒体から再生された音声信号をサンプリング周波数  $f_i$  でA/D変換するA/D変換手段と、A/D変換して得られた音声データに対して属性に基づく分析処理を行なう分析処理手段と、この分析処理で得られた情報に基づき、前記音声データを所定の時間幅を有するブロック単位に分割するブロックデータ分割手段と、分割されたブロック単位で逐次、補間処理または間引き処理を行なって、データ長を  $1/r$  倍にする一連の処理を行なって出力音声データを作成する出力音声データ作成手段と、作成された出力音声データをサンプリング周波数  $f_o$  でD/A変換して映像と同期した音声を出力する音声出力手段と、前記A/D変換するときのサンプリング周波数  $f_i$  と、前記音声データをD/A変換して音声信号にするとときのサンプリング周波数  $f_o$  と、前記速度比  $r$  とが所定の条件を満たすように、各サンプリング周波数  $f_i$  、  $f_o$  を設定するサンプリング周波数設定手段とを具備することを特徴としている。

【0014】さらに、請求項6では、請求項5に記載の記録再生装置において、前記サンプリング周波数設定手段で設定される所定の条件として、前記各サンプリング周波数  $f_i$  、  $f_o$  と、前記速度比  $r$  とが、

$$f_i / f_o = r$$

を満たすときには、

$$f_i = r f_o$$

となるサンプリング周波数  $f_i$  を使用し、また前記各サンプリング周波数  $f_i$  、  $f_o$  と、前記可変速度比  $r$  とが、

$$f_i / f_o \neq r$$

であるときには、サンプリング周波数変換係数  $c$  を含む  $f_i = r f_o / c$

となるサンプリング周波数  $f_i$  を使用することを特徴としている。

【0015】さらに、請求項7では、請求項5、または6のいずれかに記載の記録再生装置において、前記出力音声データ作成手段は、前記記録媒体から読み出される音声信号の時刻情報と前記一連の処理を経て得られた出力音声データの時刻情報を監視し、前記音声信号と前記出力音声データとの間に発生する時間のずれを補正する手段を含むことを特徴としている。

【0016】上記各請求項の構成によれば、記録媒体に音声信号を記録したときの記録速度と異なる再生速度で、前記記録媒体に記録されている音声信号を再生する際、再生された音声信号の音程を調整する音声処理方法において、正規の再生速度に対する可変速度比  $r$  で、前記記録媒体から再生された音声信号をA/D変換して音声データにするときのサンプリング周波数  $f_i$  と、前記音声データをデジタル信号処理した後、D/A変換して音声信号にするとときのサンプリング周波数  $f_o$  と、前記速度比  $r$  とが所定の条件を満たすように、各サンプリング周波数  $f_i$  、  $f_o$  を設定する。これにより、 $1/2$ 倍速再生または2倍再生したときのように、可変速度の実用的な使用範囲内で、可変速度再生された、音程が変化した音声を取り込んで、音声の明瞭さを保ちながら、映像と高精度に同期させた状態で、 $\pm 1$ オクターブ程度の範囲で、音程を元に戻す。

【0017】また、従来から使用されている回路に対し、少しの回路を付加するだけで回路を構成できる。

#### 【0018】

【発明の実施の形態】図1は本発明による音声処理方法および記録再生装置を使用した音声付きVTR可変速度再生装置の概要を示すブロック図である。

【0019】この図に示す音声付きVTR可変速度再生装置1は、VTR可変速度再生部2と、A/D変換部3と、サンプリング周波数変換部4と、分析処理部5と、ブロックデータ分割部6と、ブロックデータ蓄積部7と、接続データ生成部8と、接続データ蓄積部9と、接続順序生成部10と、音声データ接続部11と、D/A変換部12とを備えており、VTR可変速度再生部2を制御して、録音時と異なる速度で音声信号を再生したとき、サンプリング周波数変換部4によってA/D変換時のサンプリング周波数  $f_i$  (Hz) と、D/A変換時のサンプリング周波数  $f_o$  (Hz) とをマッチングさせながら、VTR可変速度再生部2の可変速度比  $r$  と、音声の属性に基づき、音声データを所定の時間幅を有するブロック単位に分割して逐次補間、または間引くことにより、VTR可変速度再生部2で可変速度再生された映像信号に、音声の提示タイミングを同期させながら、音程が変化しない高品質な音声を出力する。

【0020】この際、録音時と異なる速度で再生された音声信号をA/D変換し、デジタル信号処理によって音

程の補正と、データの時間長の伸縮とを行なった後、D/A変換して映像と同期した音声を出力する一連の過程において、A/D変換時のサンプリング周波数を $f_i$ (Hz)、D/A変換時のサンプリング周波数を $f_o$ (Hz)、VTR可変速度再生部2の可変速度比を $r$ とすると、 $f_i/f_o = r$ の関係を満たすように、サンプリング周波数 $f_i$ 、 $f_o$ を設定できる場合には、 $f_i = r f_o$ (Hz)となるサンプリング周波数 $f_i$ で、音声信号のA/D変換を行なう。またサンプリング周波数 $f_i$ 、 $f_o$ が任意の値に設定できず、 $f_i/f_o \neq r$ の関係になる場合には、A/D変換部3の後段に設けられたサンプリング周波数変換部4によって、サンプリング周波数変換係数 $c = r f_o / f_i$ (Hz)でサンプリングした状態の音声データに変換する。そして、音声の属性に基づいて分析処理を施し、この分析処理で得られた情報に基づき、前記音声データを所定の時間幅を有するブロック単位に分割し、このブロック単位で、逐次補間、または間引くことにより、データの時間長を逐次、 $1/r$ 倍にする一連の処理を行なった後、サンプリング周波数 $f_o$ (Hz)でD/A変換することにより、VTR可変速度再生部2の可変速度再生で得られた映像信号に、音声の提示タイミングを同期させながら、音程が変化しない高品質な音声を出力する。

【0021】VTR可変速度再生部2は、VTR装置に記録された音声信号を任意の映像再生速度に合わせて取り出すことができ、かつ可変再生速度したときには、正規の再生速度に対する可変速度比 $r$ を電気的に出力することができるVTR装置、または前記可変速度比 $r$ を視覚的に表示する機能を持つ既存のVTR装置などによって構成されており、通常の記録速度で記録した映像信号、音声信号を指定された再生速度で再生し、これによって得られた映像信号(变速再生映像信号)をモニタ装置などに出力しながら、この映像信号とともに再生された音声信号(变速再生音声信号)と、可変速度比情報(可変速度比 $r$ を含む情報)とをA/D変換部3に供給するとともに、前記变速再生映像信号の時刻情報と、前記可変速度比情報を接続順序生成部10に供給する。なお、この音声付きVTR可変速度再生装置1に要求される機能として、音声処理の機能のみが要求されているときには、VTR可変速度再生部2として、VTR装置以外の装置、例えば6mmのテープ音声再生装置など、音声のみのメディア記録再生装置などを使用するようにしても良い。

【0022】A/D変換部3は、VTR可変速度再生部2から出力された音声信号(变速再生音声信号)を、VTR可変速度再生部2から出力された可変速度比情報(正規の再生速度に対する可変速度比 $r$ )に応じたサンプリング周波数 $f_i$ でA/D変換し、これによって得られた音声データをサンプリング周波数変換部4に供給する。

【0023】この際、A/D変換時のサンプリング周波数 $f_i$ と、D/A変換時のサンプリング周波数 $f_o$ と、可変速度比 $r$ とが $f_i/f_o = r$ の関係を満たすように、各サンプリング周波数 $f_i$ 、 $f_o$ を設定できるときには、サンプリング周波数 $f_i = r f_o$ (Hz)で、VTR可変速度再生部2から出力される音声信号(变速再生音声信号)をA/D変換して音声データを生成する。また各サンプリング周波数 $f_i$ 、 $f_o$ が固定値である場合などのように、A/D変換時のサンプリング周波数 $f_i$ と、D/A変換時のサンプリング周波数 $f_o$ と、可変速度比 $r$ とが $f_i/f_o \neq r$ であるときには、所定のサンプリング周波数 $f_i$ (Hz)で、VTR可変速度再生部2から出力される音声信号(变速再生音声信号)をA/D変換して音声データを生成し、これによって得られた音声データをサンプリング周波数変換部4に供給する。

【0024】サンプリング周波数変換部4は、A/D変換部3から出力される音声データが $f_i/f_o = r$ の関係を満たすようにA/D変換された音声データであるときには、サンプリング周波数変換係数 $c$ として、 $c = 1$ を生成する。またA/D変換部3から出力される音声データが $f_i/f_o \neq r$ の関係でA/D変換された音声データであるときには、サンプリング周波数変換係数 $c$ として、 $c = r f_o / f_i$ を生成する。これらサンプリング周波数変換係数 $c$ に基づき、A/D変換部3から出力される音声データのサンプリング周波数を変換して、サンプリング周波数 $f_i$ でA/D変換された音声データを、正規の再生速度で再生したときと同じサンプリング周波数 $f_o$ (Hz)でA/D変換したのに相当する音声データに変換し、これを分析処理部5と、ブロックデータ分割部6とに供給する。

【0025】分析処理部5は、サンプリング周波数変換部4から出力される音声データ(サンプリング周波数 $f_o$ (Hz)でA/D変換したのに相当する音声データ)を取り込む入力処理、この入力処理で得られた音声データのサンプリングレートを下げて(本実施の形態では、4kHz)、以降の処理量を低減させるデシメーション処理、A/D変換部3から出力される音声データおよび前記デシメーション処理で得られた音声データを分析して、有声音、無声音、無音に区分する属性分析処理、この属性分析処理で得られた有声音、無声音、無音毎に、自己相関分析を行なって周期性を検出し、この検出結果に基づき、音声データを分割するのに必要なブロック長を決定するブロック長決定処理を順次行ない、これによって得られた分割情報(有声音、無声音、無音毎のブロック長を示す情報)をブロックデータ分割部6に供給する。

【0026】この場合、前記属性分析処理では、サンプリング周波数変換部4から出力される音声データについて、30ms前後の窓幅を使用して、データの自乗和を

計算し、5 m s 前後の間隔で、音声データのパワー値 P を算出するとともに、このパワー値 P と、予め設定されているしきい値  $P_{min}$  とを比較し、「 $P < P_{min}$ 」を満たす部分を無音区間と判定し、「 $P_{min} \leq P$ 」を満たす部分を有声音区間と判別する。この後、サンプリング周波数変換部 4 から出力される音声データに対して、ゼロ交差分析、前記デシメーション処理で得られた音声データに対する自己相関分析などを行ない、これらの各分析結果と、パワー値 P とに基づき、音声データのうち、

「 $P_{min} \leq P$ 」を満たす部分が声帯の振動を伴う音声区間（有声音区間）であるか、声帯の振動を伴わない音声区間（無声音区間）であるかを判定する。なお、サンプリング周波数変換部 4 から出力される音声データの各属性として、雑音や音楽などの背景音という属性も考えられるが、一般的には、雑音や背景音の信号と、音声信号とを正しく自動判別することが困難なことから、雑音、背景音も有声音、無声音のいずれかに分類される。

【0027】また、前記ブロック長決定処理では、前記属性分析処理で有声音区間と判定されたデータについて、有声音のピッチ周期が分布している 1. 25 m s ~ 2.8. 0 m s 程度の広い範囲にわたり、長さが異なる複数の窓幅を使用して自己相関分析を行なって、できるだけ正確なピッチ周期（声帯の振動周期である基本周期）を検出し、この検出結果に基づき、各ピッチ周期が各々のブロック長となるようにブロック長を決定する。また前記属性分析処理で無声音区間、無音区間とされたものについては、1. 0 m s 以内の周期性を検出し、この検出結果に基づき、ブロック長を決定し、これら有声音区間、無声音区間、無音区間の各ブロック長を分割情報としてブロックデータ分割部 6 に供給する。

【0028】ブロックデータ分割部 6 は、分析処理部 5 から出力される分割情報で示される有声音区間のブロック長、無声音区間のブロック長に基づき、サンプリング周波数変換部 4 から出力される音声データを分割し、この分割処理によって得られたブロック単位の音声データと、この音声データのブロック長とをブロックデータ蓄積部 7 と、接続データ生成部 8 とに供給する。

【0029】ブロックデータ蓄積部 7 は、リングバッファを備えており、ブロックデータ分割部 6 から出力される音声データ（ブロック単位の音声データ）と、この音声データのブロック長とを取り込み、これを前記リングバッファに一時記憶しながら、一時記憶している各ブロック長を適宜、読み出し、これを接続順序生成部 10 に供給するとともに、一時記憶しているブロック単位の音声データを適宜、読み出し、これを音声データ接続部 1 に供給する。

【0030】また、接続データ生成部 8 は、ブロックデータ分割部 6 から出力されるブロック単位の音声データを取り込みながら、図 2 に示すように、直線的に変化する窓 A、B を使用して、当該ブロックに含まれる音声デ

ータの開始部分、当該ブロックの直後または 2 つ後にあるブロックに対応する音声データの開始部分とを切り出した後、これらを重複加算して接続データを生成し、これを接続データ蓄積部 9 に供給する。

【0031】この場合、例えば、音声付き VTR 可変速度再生装置 1 の再生速度が「ゆっくり」に設定されていれば、ブロックデータ分割部 6 から出力されるブロック単位の音声データを取り込み、各ブロック毎に、当該ブロックの開始部分にある音声データに対し、時間長 d

10 (m s) の間に直線的に変化する窓 A を使用した窓掛けを行なうとともに、当該ブロック直後にあるブロックの開始部分にある音声データに対し、時間長 d (m s) の間に直線的に変化する窓 B を使用した窓掛けを行なった後、直後ブロックの開始部分と、当該ブロックの開始部分とを重複加算して、時間長 d (m s) の接続データを生成し、これを接続データ蓄積部 9 に供給する。また、音声付き VTR 可変速度再生装置 1 の再生速度が「速く」に設定されていれば、ブロックデータ分割部 6 から出力されるブロック単位の音声データを取り込み、各ブロック毎に、当該ブロックの開始部分にある音声データに対し、時間長 d (m s) の間に直線的に変化する窓 A を使用した窓掛けを行なうとともに、当該ブロックの 2 つ後にあるブロックの開始部分にある音声データに対し、時間長 d (m s) の間に直線的に変化する窓 B を使用した窓掛けを行なった後、直後ブロックの開始部分と、当該ブロックの開始部分とを重複加算して、時間長 d (m s) の接続データを生成し、これを接続データ蓄積部 9 に供給する。この際、時間長 d としては、「5 (m s)」～「当該ブロック、直後のブロック、2 つ後

30 にあるブロックの各ブロック長のうち、短い方」の範囲内にある値を任意に選択することができるが、時間長 d を短くした方が接続データ蓄積部 9 のバッファ容量を少なくすることができる。

【0032】接続データ蓄積部 9 は、リングバッファを備えており、接続データ生成部 8 から出力される接続データを取り込み、これを前記リングバッファに一時記憶しながら、一時記憶している各接続データを適宜、読み出し、これを音声データ接続部 11 に供給する。

【0033】また、接続順序生成部 10 は、VTR 可変速度再生部 2 から出力された可変速度比情報と变速再生映像の時間情報に基づき、变速再生映像の出力タイミングと音声の出力タイミングとを一致させるのに必要な時間的な伸縮倍率を算出し、これを一時記憶する書き換え可能なメモリと、予め設定されている所定の時間間隔、例えば 5.0 ~ 10.0 m s 前後の時間間隔で、前記書き換え可能なメモリに記憶されている時間的な伸縮倍率を読み出すとともに、この伸縮倍率、ブロックデータ蓄積部 7 から出力される各ブロック長、音声データ接続部 11 から出力される既接続情報に基づき、ブロック単位の各音声データと、ブロック単位の各接続データとの接

続順序、すなわち VTR 可変速度再生部 2 から出力された可変速度情報で指定された变速再生映像の出力タイミングと音声の出力タイミングとを一致させるのに必要な接続順序を時々刻々、生成する接続順序決定部とを備えている。

【0034】この場合、接続順序生成部 10 の前記接続順序決定部において、变速再生映像の出力タイミングと、音声の出力タイミングとを一致させる方法としては、VTR 可変速度再生部 2 から出力された可変速度比情報で示される可変速度比  $r$  の逆倍数  $1/r$  を計算し、これを分析処理部 5 で得られた分割情報（分割されたブロックの伸縮倍率）とし、この伸縮倍率に基づき、時間的な伸縮操作を行ない、VTR 可変速度再生部 2 から出力される变速再生映像に対し、高精度に同期させるために、音声データ接続部 11 から出力される既接続情報と、VTR 可変速度再生部 2 から出力される变速再生映像の時刻情報（データ長や CTL などの情報）を監視し、これらのずれが積算されないように、伸縮倍率を適宜、修正するという方法を使用する。

【0035】また、接続順序生成部 10 の前記接続順序決定部において、時間情報監視を行なう方法としては、図 3 に示すように、入力データ長監視部 21 によってブロックデータ蓄積部 7 から出力されるブロック単位の長さを常に計算して、入力データ長  $leng\_in$  を求めながら、出力データ目標長演算部 22 によって前記入力データ長  $leng\_in$  と、VTR 可変速度再生部 2 から出力される可変速度比  $r$  とに基づき、伸縮倍率（逆倍率） $1/r$  を計算して出力データ目標長  $target$  ( $target = leng\_in / r$ ) を求めるとともに、出力データ長監視部 23 によって音声データ接続部 11 から出力される既接続情報に基づき、出力データ長  $leng\_out$  を求めた後、比較部 24 によってこれら出力データ目標長  $target$  と、出力データ長  $leng\_out$  との差  $t$  ( $t = target - leng\_out$ ) を求め、これを音声データ接続部 11 に供給するという方法を使用する。

【0036】さらに、接続順序生成部 10 の前記接続順序決定部において、接続順序の生成方法としては、図 4、図 5 の模式図に示す方法を使用する。

【0037】まず、音声付き VTR 可変速度再生装置 1 が操作されて、再生速度が「ゆっくり」に設定されている場合には、有声音区間、無声音区間、無音区間が順次、入れ替わって出現する音声信号が入力されている状態で、音声データ接続部 11 から出力される既接続情報に基づき、音声データの属性が入れ替わったことが検知されたり、VTR 可変速度再生部 2 の再生速度が変更されて、同じ属性であっても、音声データの伸縮倍率が変更されていることが検知されたとき、接続順序の生成工程の開始条件が整ったと判定し、図 4 の (A) に示すように、このときの時刻を開始時刻  $T_0$  に決定する。

【0038】この後、この開始時刻  $T_0$  を基準として、ブロックデータ蓄積部 7 から音声データ接続部 11 に対し、既に出力された音声データのブロック長を全て加算した入力総和  $S_i$  を求めるとともに、図 4 の (A) に示すように、既に接続された音声データのブロック長を全て加算した出力総和  $S_o$  を求めた後、これら入力総和  $S_i$  と、出力総和  $S_o$  と、目的の伸縮倍率  $R$  ( $R = 1/r$  : 但し、 $R \geq 1.0$ ) と、最後に接続された音声データのブロック長  $L$  とが次の (1) 式に示す条件を満たす

10 ようなタイミングで、

$$L/2 < R \cdot S_i - S_o \quad \dots (1)$$

接続データ蓄積部 9 から出力される接続データのうち、最後に接続されたブロックに対応する接続データを置換、挿入した後、最後に接続されたブロックのうち、接続データの生成に用いた部分より後ろの部分を再度、繰り返して接続し、このブロックデータの後に、残りのブロックを順次、接続することを示す接続順序を生成し、これを音声データ接続部 11 に供給する。

【0039】これにより、図 4 の (A)、(B) に示す例では、1つ目のブロックから8つ目のブロックまでを順次、接続した時点で、前記 (1) 式に示す条件が満たされることから、8つ目のブロックの後ろに、このブロック（8つ目のブロック）に対応する接続データを置換、挿入して、8つ目のブロックのうち、接続データの生成に用いた部分より後ろの部分を繰り返し接続する。なお、この図 4 に示す例では、4つ目のブロックが既に一度、繰り返し接続されている。

【0040】また、音声付き VTR 可変速度再生装置 1 が操作されて、再生速度が「速く」に設定されている場合には、次の (2) 式に示す条件が成立立つタイミングで、

$$L/2 < S_o - R \cdot S_i \quad \dots (2)$$

接続データ蓄積部 9 から出力される接続データのうち、最後に接続されるべきブロックに対応する接続データを間引き、その直後のブロックにおいて、当該間引かれたブロックの直前と直後の接続データの生成に用いた部分より後ろの部分を接続し、このブロックの後ろに、残りのブロックを順次、接続することを示す接続順序を生成し、これを音声データ接続部 11 に供給する。

【0041】これにより、図 5 の (A)、(B) に示す接続例では、1つ目のブロックから8つ目のブロックまでを順次、接続した時点で、(2) 式に示す条件が満たされることから、8つ目のブロックの次に接続されるべき9つ目のブロックが間引かれて、10つ目のブロックのうち、接続データの生成に用いた部分より後ろの部分が接続される。なお、この図 5 に示す例では、4つ目のブロックが既に一度、間引かれている。

【0042】また、音声データ接続部 11 は、既に接続した音声データなどの接続内容（既接続情報）を接続順序生成部 10 に供給しながら、この接続順序生成部 10

から出力される差  $t$  がゼロになるように伸縮倍率を適応的に修正しながら、接続順序生成部 10 から出力される接続順序に基づき、ブロックデータ蓄積部 7 から出力されるブロック単位の音声データと、接続データ蓄積部 9 から出力される接続データとを接続し、これによって得られた一連の音声データを一時的にバッファリングしながら、D/A 変換部 12 に供給する。

【0043】D/A 変換部 12 は、音声データを記憶して FIFO 形式で出力するメモリと、所定のサンプリングレートで前記メモリから音声データを読み出し、これを D/A 変換して音声信号を生成する D/A 変換回路を備えており、音声データ接続部 11 から出力される一連の音声データを取り込んで、これをバッファリングしながら D/A 変換して音声信号を生成し、これを出力する。

【0044】このように、この実施の形態では、VTR 可変速度再生部 2 を制御して、録音時と異なる速度で音声信号を再生したとき、サンプリング周波数変換部 4 によって A/D 変換時のサンプリング周波数  $f_i$  (Hz) と、D/A 変換時のサンプリング周波数  $f_o$  (Hz) とをマッチングさせながら、VTR 可変速度再生部 2 の可変速度比  $r$  と、音声の属性に基づき、音声データを所定の時間幅を有するブロック単位に分割して逐次補間、または間引くことにより、VTR 可変速度再生部 2 で可変速度再生された映像信号に、音声の提示タイミングを同期させながら、音程が変化しない高品質な音声を出力させるようにしている。このため、1/2 倍速再生または 2 倍再生したときのように、可変速度の実用的な使用範囲内で、可変速度再生された、音程が変化した音声を取り込んで、音声の明瞭さを保ちながら、映像と高精度に同期させた状態で、音声の接続部分を最適化させつつ、±1 オクターブ程度の範囲で、音程を元に戻すことができる（請求項 1、5 の効果）。

【0045】また、この実施の形態では、正規の再生速度に対する可変速度比  $r$  で、前記記録媒体から再生された音声信号をサンプリング周波数  $f_i$  で、A/D 変換して得られた音声データをデジタル信号処理する際、前記音声データの属性を分析処理し、次にこの分析処理で得られた情報に基づき、前記音声データを所定の時間幅を有するブロック単位に分割し、ブロック単位で逐次、補間処理または間引き処理を行なって、データ長を逐次、 $1/r$  倍にする一連の処理を行なって音声データを作成した後、前記サンプリング周波数  $f_i$  で前記出力音声データを D/A 変換して音声信号を生成するようにしている。このため、1/2 倍速再生または 2 倍再生したときのように、可変速度の実用的な使用範囲内で、可変速度再生された、音程が変化した音声を取り込んで、音声の明瞭さを保ちながら、映像と高精度に同期させた状態で、音声の接続部分を最適化させつつ、±1 オクターブ程度の範囲で、音程を元に戻すことができる（請求項 2、5 の効果）。

【0046】さらに、この実施の形態では、VTR 可変速度再生部 2 から出力される音声信号の可変速度比  $r$  と、この音声信号を A/D 変換するときに使用するサンプリング周波数  $f_i$  と、音声データを D/A 変換するときに使用するサンプリング周波数  $f_o$  とが、

$$f_i / f_o = r$$

を満たすときには、

$$f_i = r f_o$$

となるサンプリング周波数  $f_i$  を使用し、また前記各サンプリング周波数  $f_i$ 、 $f_o$  と、前記可変速度比  $r$  とが、

$$f_i / f_o \neq r$$

であるときには、サンプリング周波数変換係数  $c$  を含む、

$$f_i = r f_o / c$$

となるサンプリング周波数  $f_i$  を使用するようにしている。A/D 変換部 3 と、ブロックデータ分割部 6（分析処理部 5）との間に、サンプリング周波数変換部 4 を挿入するという、従来から使用されている回路に対し、少しの回路を付加するだけで、1/2 倍速再生または 2 倍再生したときのように、可変速度の実用的な使用範囲内で、可変速度再生された、音程が変化した音声を取り込んで、音声の明瞭さを保ちながら、映像と高精度に同期させた状態で、音声の接続部分を最適化させつつ、±1 オクターブ程度の範囲で、音程を元に戻すことができる（請求項 3、5 の効果）。

【0047】さらに、この実施の形態では、VTR 可変速度再生部 2 から出力される音声信号をサンプリング周波数  $f_i$  で、A/D 変換して得られた音声データをデジタル信号処理する際、接続順序生成部 10 を構成する入力データ長監視部 21 によってブロックデータ蓄積部 7 から出力されるブロック単位の長さを常に計算して、入力データ長  $length\_in$  を求めながら、出力データ目標長演算部 22 によって前記入力データ長  $length\_in$  と、VTR 可変速度再生部 2 から出力される可変速度比  $r$  とに基づき、伸縮倍率（逆倍率） $1/r$  を計算して出力データ目標長  $target$  ( $target = length\_in / r$ ) を求めるとともに、出力データ長監視部 23 によって音声データ接続部 11 から出力される既接続情報に基づき、出力データ長  $length\_out$  を求めた後、比較部 24 によってこれら出力データ目標長  $target$  と、出力データ長  $length\_out$  との差  $t$  ( $t = target - length\_out$ ) を求め、これを音声データ接続部 11 に供給して、この差  $t$  をゼロにするように伸縮倍率を調整するようしている。このため、1/2 倍速再生または 2 倍再生したときのように、可変速度の実用的な使用範囲内で、可変速度再生された、音程が変化した音声を取り込んで、音声の明瞭さを保ちながら、音声と映像とのずれを完全に補正して、映像と高精度に同期させた状態で、音声の接続部分を最適

40

40

50

化させつつ、±1オクターブ程度の範囲で、音程を元に戻すことができる（請求項4、5の効果）。

【0048】なお、上述した実施の形態においては、接続順序生成部10によって出力データ目標長 $t_{\text{arg e t}}$ と、出力データ長 $\text{leng\_out}$ との差 $t$  ( $t = t_{\text{arg e t}} - \text{leng\_out}$ ) を求め、音声データ接続部11において、前記差 $t$ がゼロになるように、伸縮倍率を適応的に修正制御させるよう正在しているが、VTR可変速度再生部2から出力される時刻情報を音声データ接続部11に直接、フィードバックして、映像のタイミングと、音声のタイミングとの一致がとれるように、伸縮倍率を制御するようにしても良い。

【0049】また、上述した実施の形態では、A/D変換部3で使用するサンプリング周波数 $f_i$ と、D/A変換部12のサンプリング周波数 $f_o$ と、可変速度比 $r$ とが所定の関係となるように、各サンプリング周波数 $f_i$ の周波数を決定するよう正在しているが、この際地上TV放送または衛星放送Aモードの周波数帯域幅（15kHz）を確保するために、D/A変換のサンプリング周波数 $f_o$ を32kHzに固定するようにしても良い。同様に、VTR可変速度再生装置の再生速度にかかわらず、出力音声の周波数帯域幅を一定にするためにも、D/A変換のサンプリング周波数 $f_o$ を固定することが望ましい。

【0050】さらに、上述した実施の形態では、音声部分だけに対し、伸縮処理、音程調整処理などの処理を行なっているので、本発明の主要部を使用して、業務用VTR装置の統一フォーマットによる制御用インターフェースを構成し、この制御用インターフェースを既存の業務用VTR装置に外付けして、上述した音程の復元処理を行なうようにしても良い。

【0051】これによって、現在、使用されているVTR装置の記録再生方式、例えばハイビジョン方式、NTSC方式などを使用しているVTR装置でも、またどのような機種、例えばHDD-1000、D5-VTRなどの機種でも、上述した実施の形態の効果を得ることができる。

#### 【0052】

【発明の効果】以上説明したように本発明によれば、1/2倍速再生または2倍再生したときのように、可变速

度の実用的な使用範囲内で、可变速度再生された、音程が変化した音声を取り込んで、音声の明瞭さを保ちながら、映像と高精度に同期させた状態で、音声の接続部分を最適化させつつ、かつ±1オクターブ程度の範囲で、音程を元に戻すことができる。

【0053】また、従来から使用されている回路に対し、少しの回路を付加するだけの簡単な構成で装置を実現することが可能となる。

#### 【図面の簡単な説明】

【図1】本発明による音声処理方法および記録再生装置を使用した音声付きVTR可变速度再生装置の概要を示すブロック図である。

【図2】図1に示す接続データ生成部の接続データ生成動作例を示す模式図である。

【図3】図1に示す接続順序生成部内にある時間情報監視動作部分の詳細な回路構成例を示すブロック図である。

【図4】図1に示す音声付きVTR可变速度再生装置を「ゆっくり」にしたときにおける、接続順序生成部のブロック接続指示動作例を示す模式図である。

【図5】図1に示す音声付きVTR可变速度再生装置を「速く」にしたときにおける、接続順序生成部のブロック接続指示動作例を示す模式図である。

#### 【符号の説明】

1：音声付きVTR可变速度再生装置（記録再生装置）

2：VTR可变速度再生部

3：A/D変換部

4：サンプリング周波数変換部

5：分析処理部

6：ブロックデータ分割部

7：ブロックデータ蓄積部

8：接続データ生成部

9：接続データ蓄積部

10：接続順序生成部

11：音声データ接続部

12：D/A変換部

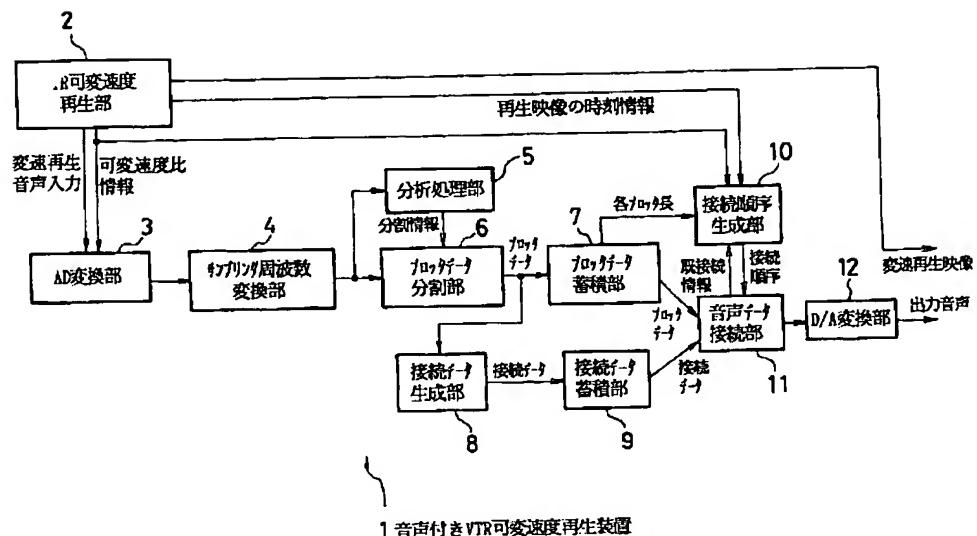
21：入力データ長監視部

22：出力データ目標長演算部

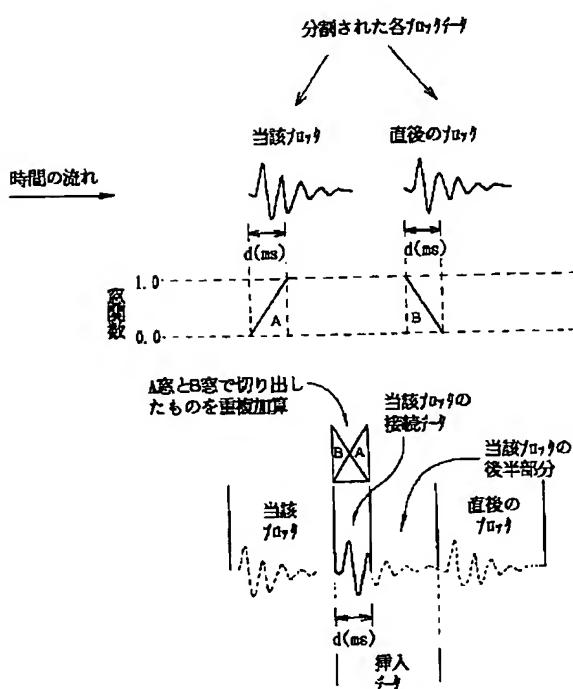
23：出力データ長監視部

24：比較部

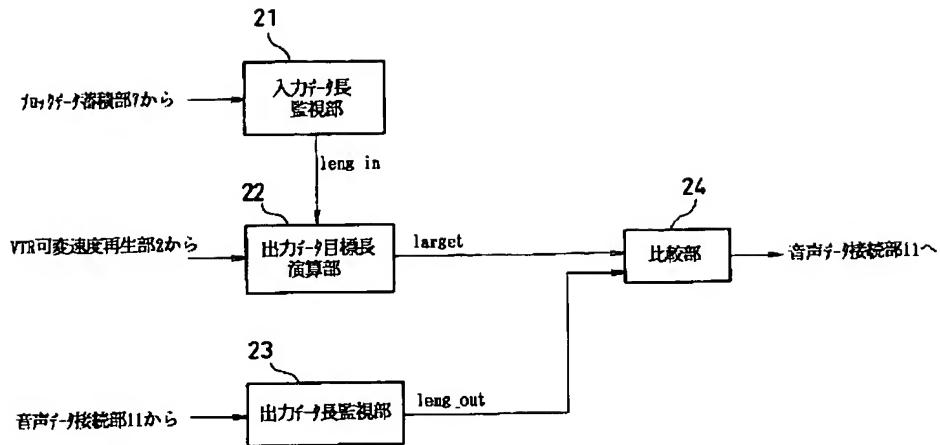
(四) 1)



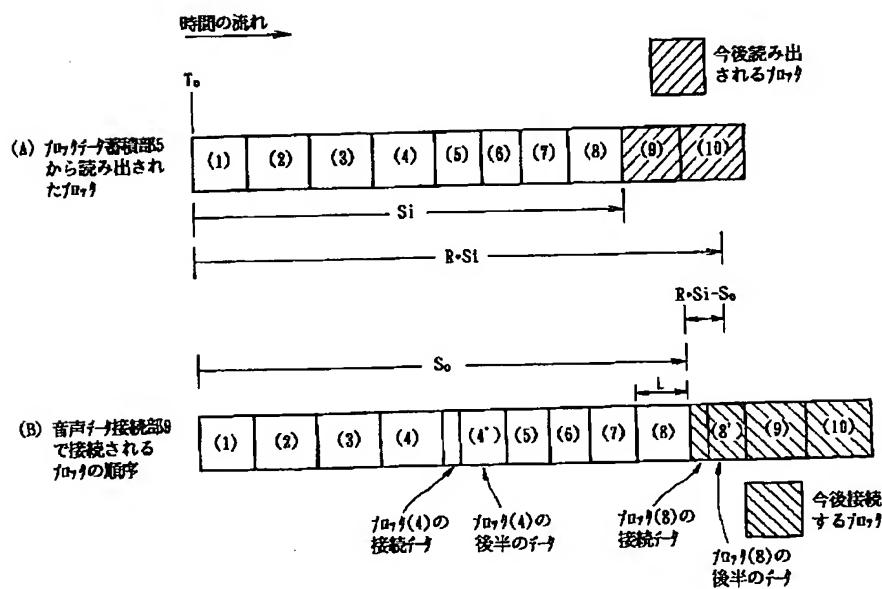
[图2]



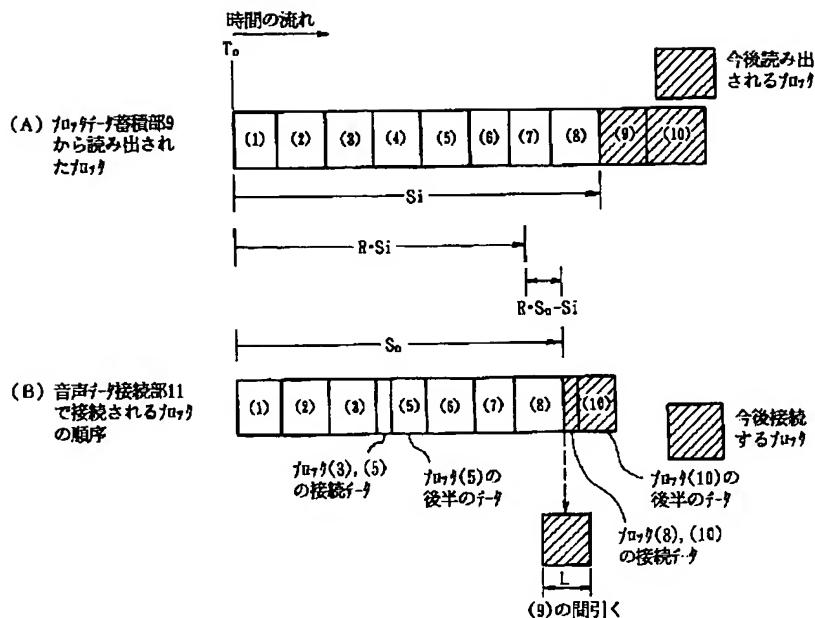
【図3】



【図4】



【図5】



## 【手続補正書】

【提出日】平成10年7月10日

## 【手続補正1】

【補正対象書類名】明細書

【補正対象項目名】発明の名称

【補正方法】変更

## 【補正内容】

【発明の名称】 音声処理方法、音声処理装置および記録再生装置

## 【手続補正2】

【補正対象書類名】明細書

【補正対象項目名】特許請求の範囲

【補正方法】変更

## 【補正内容】

## 【特許請求の範囲】

【請求項1】 記録媒体に音声信号を記録したときの速度と異なる再生速度で、前記記録媒体上の音声信号を再生する際に、再生された音声信号の音程を調整する音声処理方法において、

正規の再生速度に対する可変速度比rで、前記記録媒体から再生された音声信号をA/D変換して音声データにするときのサンプリング周波数f<sub>i</sub>と、前記音声データをデジタル信号処理した後にD/A変換して音声信号にするときのサンプリング周波数f<sub>o</sub>と、前記速度比rとが所定の条件を満たすように、各サンプリング周波数f<sub>i</sub>、f<sub>o</sub>を設定する、

ことを特徴とする音声処理方法。

【請求項2】 請求項1に記載の音声処理方法におい

て、

正規の再生速度に対する可変速度比rで、前記記録媒体から再生された音声信号をサンプリング周波数f<sub>i</sub>でA/D変換して得られた音声データをデジタル信号処理する際、

前記音声データの属性を分析処理し、次にこの分析処理で得られた情報に基づき、前記音声データを所定の時間幅を有するブロック単位に分割し、ブロック単位で逐次、補間処理または間引き処理を行なって、データ長を逐次、1/r倍にする一連の処理を行なって出力音声データを作成した後、

前記サンプリング周波数f<sub>o</sub>で前記出力音声データをD/A変換して音声信号を生成する、  
ことを特徴とする音声処理方法。【請求項3】 請求項1、または2のいずれかに記載の音声処理方法において、前記所定の条件として、前記各サンプリング周波数f<sub>i</sub>、f<sub>o</sub>と、前記速度比rとが、  
 $f_i / f_o = r$ 

を満たすときには、

$$f_i = r f_o$$

となるサンプリング周波数f<sub>i</sub>を使用し、また前記各サンプリング周波数f<sub>i</sub>、f<sub>o</sub>と、前記可変速度比rとが、  
 $f_i / f_o \neq r$ であるときには、サンプリング周波数変換係数cを含む  
 $f_i = r f_o / c$

となるサンプリング周波数  $f_i$  を使用する、ことを特徴とする音声処理方法。

【請求項4】 請求項3に記載の音声処理方法において、正規の再生速度に対する可変速度比  $r$  で、前記記録媒体から再生された音声信号をサンプリング周波数  $f_i$  で A/D 変換して得られた音声データをデジタル信号処理する際、

前記記録媒体から読み出される音声信号の時刻情報と前記一連の処理を経て得られた出力音声データの時刻情報を監視し、前記音声信号と前記出力音声データとの間に発生する時間のずれを補正する、ことを特徴とする音声処理方法。

【請求項5】 記録媒体に音声信号を記録したときの速度と異なる再生速度で、前記記録媒体上の音声信号を再生する際に、再生された音声信号の音程を調整する音声処理装置において、

前記記録媒体から再生された音声信号をサンプリング周波数  $f_i$  で A/D 変換する A/D 変換手段と、A/D 変換して得られた音声データに対して属性に基づく分析処理を行なう分析処理手段と、

この分析処理で得られた情報に基づき、前記音声データを所定の時間幅を有するブロック単位に分割するブロックデータ分割手段と、

分割されたブロック単位で逐次、補間処理または間引き処理を行なって、データ長を  $1/r$  倍にする一連の処理を行なって出力音声データを作成する出力音声データ作成手段と、

作成された出力音声データをサンプリング周波数  $f_0$  で D/A 変換して映像と同期した音声を出力する音声出力手段と、

前記 A/D 変換するときのサンプリング周波数  $f_i$  と、前記音声データを D/A 変換して音声信号にするときのサンプリング周波数  $f_0$  と、前記速度比  $r$  とが所定の条件を満たすように、各サンプリング周波数  $f_i$ 、 $f_0$  を設定するサンプリング周波数設定手段と、

を具備することを特徴とする音声処理装置。

【請求項6】 請求項5に記載の音声処理装置において、前記サンプリング周波数設定手段で設定される所定の条件として、前記各サンプリング周波数  $f_i$ 、 $f_0$  と、前記速度比  $r$  とが、

$$f_i / f_0 = r$$

を満たすときには、

$$f_i = r f_0$$

となるサンプリング周波数  $f_i$  を使用し、

また前記各サンプリング周波数  $f_i$ 、 $f_0$  と、前記可変速度比  $r$  とが、

$$f_i / f_0 \neq r$$

であるときには、サンプリング周波数変換係数  $c$  を含む

$$f_i = r f_0 / c$$

となるサンプリング周波数  $f_i$  を使用することを特徴とする音声処理装置。

【請求項7】 請求項5、または6のいずれかに記載の音声処理装置において、前記出力音声データ作成手段は、前記記録媒体から読み出される音声信号の時刻情報と前記一連の処理を経て得られた出力音声データの時刻情報を監視し、前記音声信号と前記出力音声データとの間に発生する時間のずれを補正する手段を含むことを特徴とする音声処理装置。

【請求項8】 請求項5乃至7のいずれかに記載された音声処理装置を含むことを特徴とする記録再生装置。

#### 【手続補正3】

【補正対象書類名】明細書

【補正対象項目名】0001

【補正方法】変更

【補正内容】

【0001】

【発明の属する技術分野】本発明は、業務用 VTR 装置、6mm テープレコーダ装置など、再生速度に比例して音程が変化した音声を出力する記録再生装置において、記録された音声を任意の速度で再生する際に、再生速度に対応した正規の音程を持つ音声を出力する音声処理方法および記録再生装置に係わり、特に VTR 装置に適用するとき、好みの再生速度に即応し、かつ映像に同期した高品質な音声を出力する音声処理方法、音声処理装置および記録再生装置に関する。

#### 【手続補正4】

【補正対象書類名】明細書

【補正対象項目名】0002

【補正方法】変更

【補正内容】

【0002】 【発明の概要】本発明は、業務用 VTR 装置などにおいて、スローモーション映像などを収録時と異なる速さで再生する場合、再生速度に比例して音程が変化した音声が出力されるが、この音程が変化した音声を任意のサンプリング周波数  $f_i$  で A/D 変換して得られた信号を正規の再生速度に対する可変速度比  $r$  の逆倍数  $1/r$  のサンプリング周波数  $f_i/r$  (Hz) のデジタル信号に変換するとともに、音声のタイミングを映像に同期させるため、サンプリング周波数  $f_i/r$  (Hz) にサンプリング変換した音声データを基本周期またはフレーム単位など、所定の時間幅を有するブロック単位に分割して蓄積し、VTR 装置からの入力データ時間長に等しくなるように、逐次、ブロック単位での補間、または間引き操作を行なうことにより、VTR 装置を任意の速度で再生する際、その速度の可変タイミングに即応し、かつ映像に同期して音程が変化しない音声を出力する音声処理方法と、この音声処理方法を実現する音声処理装置および記録再生装置を提供するものである。

**【手続補正5】****【補正対象書類名】**明細書**【補正対象項目名】**0008**【補正方法】**変更**【補正内容】**

**【0008】**本発明は上記の事情に鑑み、 $1/2$ 倍速再生または2倍再生したときのように、可変速度の実用的な使用範囲内で、可変速度再生された、音程が変化した音声を取り込んで、音声の明瞭さを保ちながら、映像と高精度に同期させた状態で、±1オクターブ程度の範囲で、音程を元に戻すことができる音声処理方法、音声処理装置および記録再生装置を提供することを目的としている。

**【手続補正6】****【補正対象書類名】**明細書**【補正対象項目名】**0013**【補正方法】**変更**【補正内容】**

**【0013】**さらに、請求項5では、記録媒体に音声信号を記録したときの速度と異なる再生速度で、前記記録媒体上の音声信号を再生する際に、再生された音声信号の音程を調整する音声処理装置において、前記記録媒体から再生された音声信号をサンプリング周波数  $f_i$  で A/D 変換する A/D 変換手段と、A/D 変換して得られた音声データに対して属性に基づく分析処理を行なう分析処理手段と、この分析処理で得られた情報に基づき、前記音声データを所定の時間幅を有するブロック単位に分割するブロックデータ分割手段と、分割されたブロック単位で逐次、補間処理または間引き処理を行なって、データ長を  $1/r$  倍にする一連の処理を行なって出力音声データを作成する出力音声データ作成手段と、作成された出力音声データをサンプリング周波数  $f_o$  で D/A 変換して映像と同期した音声を出力する音声出力手段と、前記 A/D 変換するときのサンプリング周波数  $f_i$  と、前記音声データを D/A 変換して音声信号にするときのサンプリング周波数  $f_o$  と、前記速度比  $r$  とが所定の条件を満たすように、各サンプリング周波数  $f_i$ 、 $f_o$  を設定するサンプリング周波数設定手段とを具備することを特徴としている。

**【手続補正7】****【補正対象書類名】**明細書**【補正対象項目名】**0014**【補正方法】**変更**【補正内容】**

**【0014】**さらに、請求項6では、請求項5に記載の音声処理装置において、前記サンプリング周波数設定手

段で設定される所定の条件として、前記各サンプリング周波数  $f_i$ 、 $f_o$  と、前記速度比  $r$  とが、

$$f_i / f_o = r$$

を満たすときには、

$$f_i = r f_o$$

となるサンプリング周波数  $f_i$  を使用し、また前記各サンプリング周波数  $f_i$ 、 $f_o$  と、前記可変速度比  $r$  とが、

$$f_i / f_o \neq r$$

であるときには、サンプリング周波数変換係数  $c$  を含む  $f_i = r f_o / c$

となるサンプリング周波数  $f_i$  を使用することを特徴としている。

**【手続補正8】****【補正対象書類名】**明細書**【補正対象項目名】**0015**【補正方法】**変更**【補正内容】**

**【0015】**さらに、請求項7では、請求項5、または6のいずれかに記載の音声処理装置において、前記出力音声データ作成手段は、前記記録媒体から読み出される音声信号の時刻情報と前記一連の処理を経て得られた出力音声データの時刻情報を監視し、前記音声信号と前記出力音声データとの間に発生する時間のずれを補正する手段を含むことを特徴としている。加えて、請求項8の記録再生装置は、請求項5乃至7のいずれかに記載された音声処理装置を含むことを特徴としている。

**【手続補正9】****【補正対象書類名】**明細書**【補正対象項目名】**0018**【補正方法】**変更**【補正内容】****【0018】**

**【発明の実施の形態】**図1は本発明による音声処理方法、音声処理装置および記録再生装置を使用した音声付きVTR可変速度再生装置の概要を示すブロック図である。

**【手続補正10】****【補正対象書類名】**明細書**【補正対象項目名】**図1**【補正方法】**変更**【補正内容】**

**【図1】**本発明による音声処理方法、音声処理装置および記録再生装置を使用した音声付きVTR可変速度再生装置の概要を示すブロック図である。